Pose Tracking II



Gordon Wetzstein Stanford University

EE 267 Virtual Reality

Lecture 12

stanford.edu/class/ee267/

WARNING

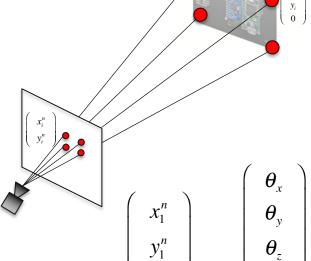
- this class will be dense!
- will learn how to use nonlinear optimization (Levenberg-Marquardt algorithm) for pose estimation
- why ???
 - more accurate than homography method
 - can dial in lens distortion estimation, and estimation of intrinsic parameters (beyond this lecture, see lecture notes)
 - LM is very common in 3D computer vision → camera-based tracking

 goal: estimate pose via nonlinear least squares optimization

$$\underset{\{p\}}{\text{minimize}} \left\| b - f(g(p)) \right\|_{2}^{2}$$

$$\underset{\text{image formation}}{\bullet}$$

- minimize reprojection error
- pose p is 6-element vector with 3
 Euler angles and translation of
 VRduino w.r.t. base station



Overview

- review: gradients, Jacobian matrix, chain rule, iterative optimization
- nonlinear optimization: Gauss-Newton, Levenberg-Marquardt
- pose estimation using LM
- pose estimation with VRduino using nonlinear optimization

Review

Review: Gradients

• gradient of a function that depends on multiple variables:

$$\frac{\partial}{\partial x} f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$$

$$f:\mathfrak{R}^n\to\mathfrak{R}$$

Review: The Jacobian Matrix

 gradient of a vector-valued function that depends on multiple variables:

$$\frac{\partial}{\partial x} f(x) = J_f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

$$f: \mathbb{R}^n \to \mathbb{R}^m, \quad J_f \in \mathbb{R}^{m \times n}$$

Review: The Chain Rule

• here's how you've probably been using it so far:

$$\frac{\partial}{\partial x} f(g(x)) = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$$

this rule applies when $f: \Re \to \Re$ $g: \Re \to \Re$

Review: The Chain Rule

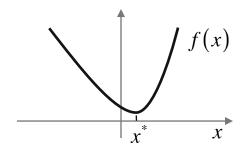
• here's how it is applied in general:

$$\frac{\partial}{\partial x} f(g(x)) = J_f \cdot J_g = \begin{pmatrix} \frac{\partial f_1}{\partial g_1} & \cdots & \frac{\partial f_1}{\partial g_o} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial g_1} & \cdots & \frac{\partial f_m}{\partial g_o} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_o}{\partial x_1} & \cdots & \frac{\partial g_o}{\partial x_n} \end{pmatrix}$$

 $f: \mathbb{R}^o \to \mathbb{R}^m, \quad g: \mathbb{R}^n \to \mathbb{R}^o, \quad J_f \in \mathbb{R}^{m \times o}, \quad J_g \in \mathbb{R}^{o \times n}$

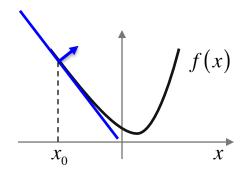
Review: Minimizing a Function

goal: find point x* that minimizes a nonlinear function f(x)



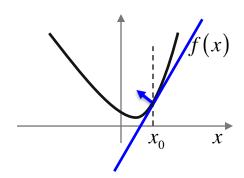
Review: What is a Gradient?

• gradient of f at some point x_0 is the slope at that point



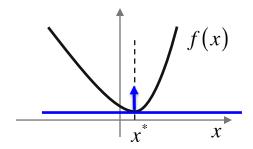
Review: What is a Gradient?

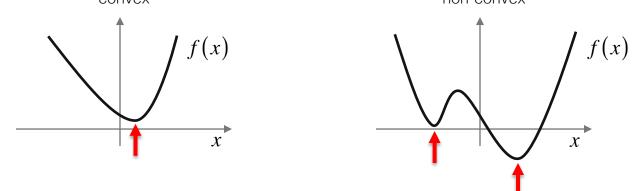
• gradient of f at some point x_0 is the slope at that point



Review: What is a Gradient?

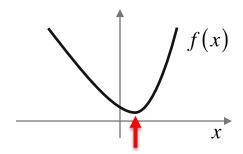
• extremum is where gradient is 0! (sometimes have to check 2nd derivative to see if it's a minimum and not a maximum or saddle point)



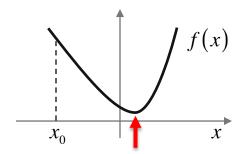


- convex optimization: there is only a single global minimum
- non-convex optimization: multiple local minima

• how to find where gradient is 0?

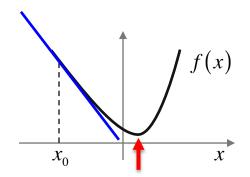


how to find where gradient is 0?



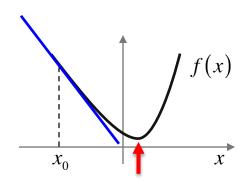
1. start with some initial guess x_0 , e.g. a random value

how to find where gradient is 0?



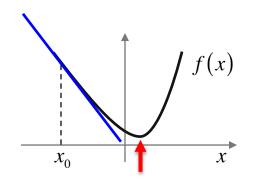
- 1. start with some initial guess x_0 , e.g. a random value
- 2. update guess by linearizing function and minimizing that

how to linearize a function? → using Taylor expansion!



$$f(x_0 + \Delta x) \approx f(x_0) + J_f \Delta x + \varepsilon$$

• find minimum of linear function approximation

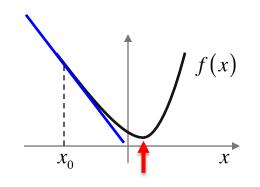


$$f(x_0 + \Delta x) \approx f(x_0) + J_f \Delta x + \varepsilon$$

minimize $\|b - f(x_0 + \Delta x)\|_2^2$

$$\approx \left\| b - \left(f(x_0) + J_f \Delta x \right) \right\|_2^2$$

find minimum of linear function approximation (gradient=0)



$$f(x_0 + \Delta x) \approx f(x_0) + J_f \Delta x + \varepsilon$$

minimize $\|b - f(x_0 + \Delta x)\|_2^2$

$$\approx \left\| b - \left(f(x_0) + J_f \Delta x \right) \right\|_2^2$$

equate gradient to zero:

$$0 = J_f^T J_f \Delta x - J_f^T (b - f(x))$$



$$\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$$
normal equations

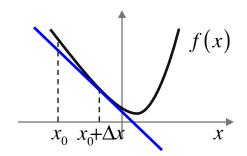
take step and repeat procedure

$$f(x)$$

$$x_0 \quad x_0 + \Delta x \quad x$$

$$\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$$

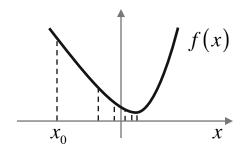
take step and repeat procedure, will get there eventually



$$\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$$

Review: Optimization – Gauss-Newton

· results in an iterative algorithm

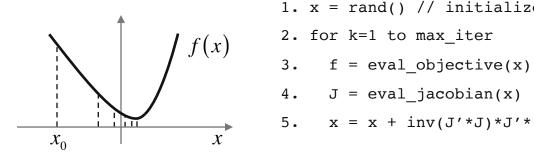


$$x_{k+1} = x_k + \Delta x$$

$$\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$$

Review: Optimization – Gauss-Newton

results in an iterative algorithm



- 1. x = rand() // initialize x0
- 2. for k=1 to max_iter

 $\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$

5. x = x + inv(J'*J)*J'*(b-f) // update x

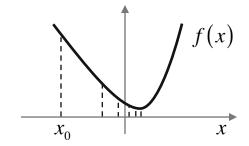
Review: Optimization - Gauss-Newton

• matrix $\mathcal{J}\mathcal{J}$ can be ill-conditioned (i.e. not invertible)

$$\Delta x = \left(J_f^T J_f\right)^{-1} J_f^T \left(b - f(x)\right)$$

Review: Optimization – Levenberg

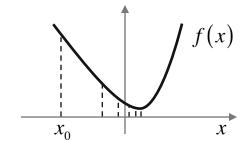
- matrix $\mathcal{J}\mathcal{J}$ can be ill-conditioned (i.e. not invertible)
- add a diagonal matrix to make invertible acts as damping



$$\Delta x = \left(J_f^T J_f + \lambda I\right)^{-1} J_f^T \left(b - f(x)\right)$$

Review: Optimization – Levenberg-Marquardt

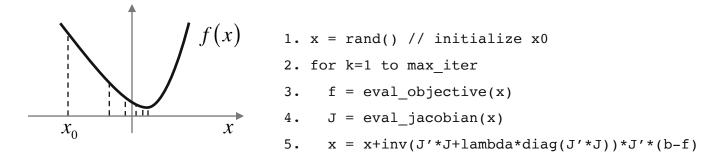
- matrix $\mathcal{J}\mathcal{J}$ can be ill-conditioned (i.e. not invertible)
- better: use $\mathcal{J}^T\mathcal{J}$ instead of /as damping. This is LM!



$$\Delta x = \left(J_f^T J_f + \lambda \operatorname{diag}\left(J_f^T J_f\right)\right)^{-1} J_f^T \left(b - f(x)\right)$$

Review: Optimization – Levenberg-Marquardt

- matrix $\mathcal{J}^T \mathcal{J}$ can be ill-conditioned (i.e. not invertible)
- better: use $\mathcal{J}\mathcal{J}$ instead of /as damping. This is LM!



$$\Delta x = \left(J_f^T J_f + \lambda \operatorname{diag}\left(J_f^T J_f\right)\right)^{-1} J_f^T \left(b - f(x)\right)$$

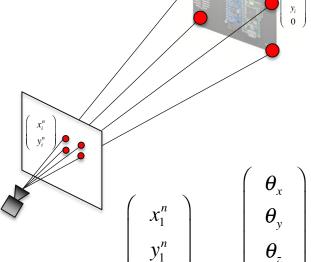
Pose Estimation via Levenberg-Marquardt

 goal: estimate pose via nonlinear least squares optimization

$$\underset{\{p\}}{\text{minimize}} \left\| b - f(g(p)) \right\|_{2}^{2}$$

$$\underset{\text{image formation}}{\bullet}$$

- minimize reprojection error
- pose p is 6-element vector with 3
 Euler angles and translation of
 VRduino w.r.t. base station



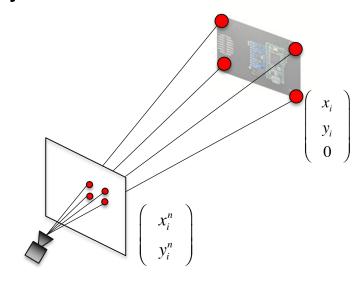
Pose Estimation - Objective Function

 goal: estimate pose via nonlinear least squares optimization

$$\underset{p}{\text{minimize}} \left\| b - f(g(p)) \right\|_{2}^{2}$$

$$\underset{\text{image formation}}{\uparrow}$$

 objective function is sum of squares of reprojection error



$$\left|\left|b-f(g(p))\right|\right|_{2}^{2} = \left(x_{1}^{n}-f_{1}(g(p))\right)^{2}+\left(y_{1}^{n}-f_{2}(g(p))\right)^{2}+\ldots+\left(x_{4}^{n}-f_{7}(g(p))\right)^{2}+\left(y_{4}^{n}-f_{8}(g(p))\right)^{2}$$

Image Formation

1. transform 3D point into view space:

$$\begin{pmatrix} x_{i}^{c} \\ y_{i}^{c} \\ w_{i}^{c} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & t_{x} \\ r_{21} & r_{22} & t_{y} \\ r_{31} & r_{32} & t_{z} \end{pmatrix} \cdot \begin{pmatrix} x_{i} \\ y_{i} \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} h_{1} & h_{2} & h_{3} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{i} \\ x_{i} \\ x_{i} \end{pmatrix}$$

$$\begin{bmatrix} y_i^c \\ w_i^c \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \cdot \begin{bmatrix} y_i \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_6 & h_6 & h_6 \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

$$\underbrace{2. \text{ perspective divide}}_{n_7} \cdot \underbrace{n_8}_{n_9} \cdot \underbrace{n_9}_{n_1} \cdot \underbrace{1}_{n_2} \cdot$$

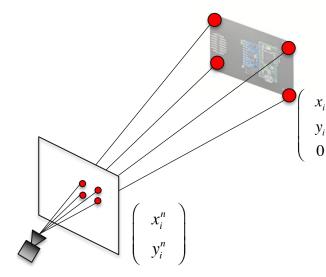


Image Formation: g(p) and f(h)

split up image formation into two functions

$$f(h) = f(g(p))$$
$$g: \Re^6 \to \Re^9, \quad f: \Re^9 \to \Re^8$$

Image Formation: f(h)

 f(h) uses elements of homography matrix h to compute projected 2D coordinates as

$$f(h) = \begin{pmatrix} f_1(h) \\ f_2(h) \\ \vdots \\ f_7(h) \\ f_8(h) \end{pmatrix} = \begin{pmatrix} x_1^n \\ y_1^n \\ \vdots \\ x_4^n \\ y_4^n \end{pmatrix} = \begin{pmatrix} \frac{h_1x_1 + h_2y_1 + h_3}{h_7x_1 + h_8y_1 + h_9} \\ \frac{h_4x_1 + h_5y_1 + h_6}{h_7x_1 + h_8y_1 + h_9} \\ \vdots \\ \frac{h_1x_4 + h_2y_4 + h_3}{h_7x_4 + h_8y_4 + h_9} \\ \frac{h_4x_4 + h_5y_4 + h_6}{h_7x_4 + h_8y_4 + h_9} \end{pmatrix}$$

Jacobian Matrix of f(h)

Jacobian Matrix of
$$f(h)$$

$$J_{f} = \begin{bmatrix} \frac{\partial f_{1}}{\partial h_{1}} & \dots & \frac{\partial f_{1}}{\partial h_{9}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{8}}{\partial h_{1}} & \dots & \frac{\partial f_{8}}{\partial h_{9}} \end{bmatrix}$$

$$f_{1}(h) = \frac{h_{1} x_{1} + h_{2} y_{1} + h_{3}}{h_{7} x_{1} + h_{9} y_{1} + h_{9}}$$

first row of Jacobian matrix



$$\frac{\partial f_{1}}{\partial h_{2}} = \frac{y_{1}}{h_{7} x_{1} + h_{8} y_{1} + h_{9}}$$

$$\frac{\partial f_{1}}{\partial h_{2}} = \frac{y_{1}}{h_{7} x_{1} + h_{8} y_{1} + h_{9}}$$

$$\frac{\partial f_{1}}{\partial h_{3}} = \frac{1}{h_{7} x_{1} + h_{8} y_{1} + h_{9}}$$

$$\frac{\partial f_{1}}{\partial h_{4}} = 0, \quad \frac{\partial f_{1}}{\partial h_{5}} = 0, \quad \frac{\partial f_{1}}{\partial h_{6}} = 0$$

$$\frac{\partial f_1}{\partial h_7} = -\left(\frac{h_1 x_1 + h_2 y_1 + h_3}{\left(h_7 x_1 + h_8 y_1 + h_9\right)^2}\right) x_1$$

$$\frac{\partial f_1}{\partial h_7} = -\left[\frac{1}{\left(h_7 x_1 + h_8 y_1 + h_9\right)^2}\right] x_1$$

$$\frac{\partial f_1}{\partial h_8} = -\left(\frac{h_1 x_1 + h_2 y_1 + h_3}{\left(h_7 x_1 + h_8 y_1 + h_9\right)^2}\right) y_1$$

 $\frac{\partial f_1}{\partial h_9} = -\frac{h_1 x_1 + h_2 y_1 + h_3}{\left(h_7 x_1 + h_8 y_1 + h_9\right)^2}$

Jacobian Matrix of f(h)

$$\left(\begin{array}{ccc} \vdots & \ddots & \vdots \\ \frac{\partial f_8}{\partial h_1} & \cdots & \frac{\partial f_8}{\partial h_9} \end{array}\right)$$

 the remaining rows of the Jacobian can be derived with a similar pattern

 see course notes for a detailed deriavtion of the elements of this Jacobian matrix

Image Formation: g(p)

q(p) uses 6 pose parameters to compute elements of homography matrix h as

$$g(p) = \begin{pmatrix} g_1(p) \\ \vdots \\ g_9(p) \end{pmatrix} = \begin{pmatrix} h_1 \\ \vdots \\ h_9 \end{pmatrix}$$

rotation matrix from Euler angles:

$$R = R_1(\theta_1) \cdot R_2(\theta_2) \cdot R_2(\theta_3)$$

$$R = R_z(\theta_z) \cdot R_x(\theta_x) \cdot R_y(\theta_y)$$

$$r_{13} \atop r_{23} = \begin{pmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \end{pmatrix} \begin{pmatrix} \cos(\theta_y) & 0 & \sin(\theta_z) \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \end{pmatrix}$$

Image Formation: g(p)

write as

$$h_1 = g_1(p) = \cos(\theta_y)\cos(\theta_z) - \sin(\theta_x)\sin(\theta_y)\sin(\theta_z)$$

$$h_2 = g_2(p) = -\cos(\theta_x)\sin(\theta_z)$$

$$h_3 = g_3(p) = t_x$$

$$h_4 = g_4(p) = \cos(\theta_y)\sin(\theta_z) + \sin(\theta_x)\sin(\theta_y)\cos(\theta_z)$$

$$h_5 = g_5(p) = \cos(\theta_x)\cos(\theta_z)$$

$$h_6 = g_6(p) = t_y$$

$$h_7 = g_7(p) = \cos(\theta_x)\sin(\theta_y)$$

$$h_8 = g_8(p) = -\sin(\theta_x)$$

$$h_9 = g_9(p) = -t_z$$

Jacobian Matrix of g(p)

$$p = (p_1, p_2, p_3, p_4, p_5, p_6) = (\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$$

$$= \left(\begin{array}{ccc} \partial p_1 & \partial p_6 \\ \vdots & \ddots & \vdots \\ \frac{\partial g_9}{\partial p_1} & \cdots & \frac{\partial g_9}{\partial p_6} \end{array}\right)$$

$$h_2 = g_2(p) = -\cos(\theta_x)\sin(\theta_z)$$

$$h_3 = g_3(p) = t_x$$

$$h_4 = g_4(p) = \cos(\theta_y)\sin(\theta_z) + \sin(\theta_x)\sin(\theta_y)\cos(\theta_z)$$

$$h_5 = g_5(p) = \cos(\theta_x)\cos(\theta_z)$$

$$h_6 = g_6(p) = t_y$$

$$h_7 = g_7(p) = \cos(\theta_x)\sin(\theta_y)$$

$$h_8 = g_8(p) = -\sin(\theta_x)$$

 $h_0 = g_0(p) = -t_z$

 $h_1 = g_1(p) = \cos(\theta_y)\cos(\theta_z) - \sin(\theta_x)\sin(\theta_y)\sin(\theta_z)$

Jacobian Matrix of g(p)

Jacobian Matrix of
$$g(p)$$

 $g(p) = (\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$

$$p = (p_1, p_2, p_3, p_4, p_5, p_6) = (\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$$

$$J_g = \begin{bmatrix} \frac{\partial g_9}{\partial p_1} & \dots & \frac{\partial g_9}{\partial p_6} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_9}{\partial p_1} & \dots & \frac{\partial g_9}{\partial p_6} \end{bmatrix}$$

$$h_1 = g_1(p) = \cos(\theta_y)\cos(\theta_z) - \sin(\theta_x)\sin(\theta_y)\sin(\theta_z)$$

$$\frac{\partial g_1}{\partial p_1} = -\cos(\theta_x)\sin(\theta_y)\sin(\theta_z)$$

$$\frac{\partial g_1}{\partial p_2} = -\sin(\theta_y)\cos(\theta_z) - \sin(\theta_x)\cos(\theta_y)\sin(\theta_z)$$

$$\frac{\partial g_1}{\partial p_3} = -\cos(\theta_y)\sin(\theta_z) - \sin(\theta_x)\sin(\theta_y)\cos(\theta_z)$$

$$\frac{\partial g_1}{\partial p_4} = 0, \quad \frac{\partial g_1}{\partial p_5} = 0, \quad \frac{\partial g_1}{\partial p_6} = 0$$

Jacobian Matrix of g(p)

$$= \begin{bmatrix} \frac{\partial g_1}{\partial p_1} & \dots & \frac{\partial g_1}{\partial p_6} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_9}{\partial p_1} & \dots & \frac{\partial g_9}{\partial p_6} \end{bmatrix}$$

 the remaining rows of the Jacobian can be derived with a similar pattern

 see course notes for a detailed deriavtion of the elements of this Jacobian matrix

Jacobian Matrices of f and g

• to get the Jacobian of f(g(p)), compute the two Jacobian matrices and multiply them

$$J = J_f \cdot J_g = \begin{pmatrix} \frac{\partial f_1}{\partial h_1} & \cdots & \frac{\partial f_1}{\partial h_9} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_8}{\partial h_1} & \cdots & \frac{\partial f_8}{\partial h_9} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial p_1} & \cdots & \frac{\partial g_1}{\partial p_6} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_9}{\partial p_1} & \cdots & \frac{\partial g_9}{\partial p_6} \end{pmatrix}$$

Pose Tracking with LM

LM then iteratively updates pose as

$$p^{(k+1)} = p^{(k)} + \left(J^T J + \lambda \operatorname{diag}\left(J^T J\right)\right)^{-1} J^T \left(b - f\left(g\left(p^{(k)}\right)\right)\right)$$

PSeudo-code 1. p = ... // initialize p0
2. for k=1 to max_iter
3. f = eval_objective(p)
4. J = get_jacobian(p)
5. p = p + inv(J'*J+lambda*diag(J'*J))*J'*(b-f)

Pose Tracking with LM

- 1. value = function eval objective(p)
- for i=1:4

- value(2*(i-1)) = ...
- 3.
- $\frac{h_1 x_i + h_2 y_i + h_3}{h_7 x_i + h_8 y_i + h_9}$ value(2*(i-1)+1) = ...4.
 - $h_4 x_i + h_5 y_i + h_6$ $h_7 x_i + h_8 y_i + h_9$

Pose Tracking with LM

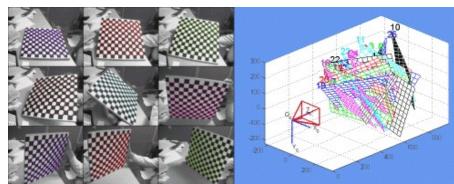
$$J_{f} = \begin{pmatrix} \frac{\partial f_{1}}{\partial h_{1}} & \dots & \frac{\partial f_{1}}{\partial h_{9}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{8}}{\partial h_{1}} & \dots & \frac{\partial f_{8}}{\partial h_{9}} \end{pmatrix}$$
2. Jf = get_jacobian_f(g(p))
3. Jg = get_jacobian_g(p)
$$J_{g} = \begin{pmatrix} \frac{\partial g_{1}}{\partial p_{1}} & \dots & \frac{\partial g_{1}}{\partial p_{6}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{9}}{\partial p_{1}} & \dots & \frac{\partial g_{9}}{\partial p_{6}} \end{pmatrix}$$

Pose Tracking with LM on VRduino

- some more hints for implementation:
 - let Arduino Matrix library compute matrix-matrix multiplications and also matrix inverses for you!
 - run homography method and use that to initialize p for LM
 - use something like 5-25 interations of LM per frame for real-time performance
 - user-defined parameter λ
 - good luck!

Outlook: Camera Calibration

- camera calibration is one of the most fundamental problems in computer vision and imaging
- task: estimate intrinsic (lens distortion, focal length, principle point) & extrinsic (translation, rotation) camera parameters given images of planar checkerboards
- uses similar procedure as discussed today



http://www.vision.caltech.edu/bouguetj/calib_doc/

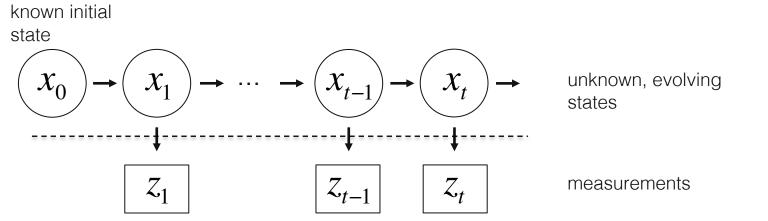
Outlook: Sensor Fusion with Extended Kalman Filter

- also desirable: estimate bias of each of all IMU sensors
- also desirable: joint pose estimation from all IMU + photodiode measurements

 can do all of that with an Extended Kalman Filter - slightly too advanced for this class, but you can find a lot of literature in the robotic vision community

Outlook: Sensor Fusion with Extended Kalman Filter

- Extended Kalman filter: can be interpreted as a Bayesian framework for sensor fusion
- Hidden Markov Model (HMM)



Must read: course notes on tracking!